本周工作：

1. 修改教科书：

   增加了几个软件的案例及其说明。把大部分的软件说明进行了扩充。对文字进行了修订。重要的软件插入图片。

2. 轨迹查询项目：

   修改 introduction，撰写 related work。主要修订文字的表达和 introduction 后半部分。并添加相关引用。Relatedwork 主要分为轨迹的可视化和使用 NLP 的可视查询系统。还是以讲趋势的方法来写。

   程序方面：

   高完成了后台的查询接口，正在训练数据

   俞完成了之前说的时间轴上的编码，现在在优化选择，接入系统中。

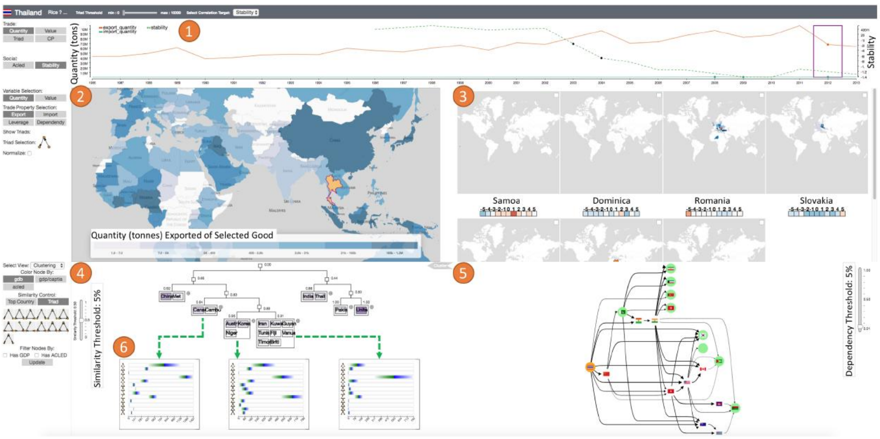
   没来的徐远程处理出租车轨迹数据，并接入之前她写的轨迹聚类方法。调试中。

3. 准备下周组会报告

   读论文，做 PPT

4. 文章阅读：

《A Visual Analytics Framework for Spatiotemporal Trade Network Analysis》



对国际贸易关系进行可视化，节和国际事件，分析贸易关系。数据为大规模多维度的节点连接图数据。

相关工作写作中读的文章：

和分析城市区域语义相关文章：

《Applications of Trajectory Data From the Perspective of a Road Transportation Agency: Literature Review and Maryland Case Study》

《Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns》使用的数据也是手机基站数据，另外结合了社交网络数据。

《The livehoods project: Utilizing social media to understand the dynamics of a city》

《Exploiting Semantic Annotations for Clustering Geographic Areas and Users in Location-based Social Networks. 》

《Visualizing interchange patterns in massive movement data》

轨迹可视化相关文章：

《Visual Abstraction of Large Scale Geospatial Origin-Destination Movement Data》

大规模 OD 轨迹数据的可视化方法。

Introduction 进展：

Querying is a common task when analyzing massive trajectory data. To do this, data analysts present their requirements by inputting query conditions to get a group of filtered trajectories. Such query mechanism has proven to be useful for filtering data which can be applied to make improvement in peoples life quality [32] [cite here], urban environment [cite here], and city operation systems [48, 54]. However, the task becomes complex when analysts are from different domains and have different query requirements. A crucial problem is how to help the analysts express their query requirements. Many research worked on helping the analysts define spatial-temporal conditions by means of an interactive interface [cite here] [10, 58]. Using textual sentence is one of the most natural ways to express complicated query conditions. By externalizing the geospatial locations with inherent contextual information (a.k.a.,

textualization), annotating trajectory data provides knowledge-based context information [46] which dramatically enriches the raw data [2, 20, 39, 52]. In a textual sentence, the analysts use location names (e.g., Golden Gate Bridge), functions categories (e.g., Education areas, residential areas) or directions to demonstrate the spatial constraints. These constraints can be further applied to annotated trajectory data to filter the massive trajectories.

Despite the "textualization-and-query" scheme has been applied in several visual analytics systems including studying taxi trajectories [1, 11] and human mobility patterns [55], most existing methods assume the input trajectories with accurate geo-location. A trajectory with inaccurate geo-location is represented by a series of chronologically ordered sampling points. Each point only tells us a spatial region that the moving object may be located in, but not the accurate longitude

and latitude. Annotating such inaccurate trajectory faces a challenge on assigning a sampling point to one fixed nominal value of context (e.g., a street or POI name). For example, if a trajectory enters a region which contains a restaurant and a company, then either the restaurant or the company is likely to be the destination (potential context) of that trajectory. Although it is difficult to confirm a unique destination in an inaccurate trajectory, integrating high-level description of regions (e.g., functions) and temporal attributes into the textualization process can help to speculate it. For example, people are more likely to go to a restaurant during the lurch time than a school. During the weekend, It is more likely for a person to go shopping in a mall than work in a

company in an entertainment area. Recently, we have witnessed the success of many approaches to extract the high-level description of regions [13,34,44,47]. Yuan et al. [52] discover the regions' functions by using a topic-based inference model on Points of interests (POI) data. We integrate their topic-based method into a novel textualization process. In the process, we take descriptive information and temporal attributes into account to achieve a reasonable trajectory annotating method. [I still work on the following sentence]One other issue is how to speed up the query. With a large volume of data (e.g., 14 million citizens' trajectories of multiple months with more than 100 GB), A real-time visual analytic system demands efficient indexing mechanism.

The issue of inputting query condition becomes even more difficult when the analysts don't have a clearly retrieved goal. For example, when a manager of a university wants to analyze the daily life of students, it is hard to express what is the 'daily life'. The analysts only have some fuzzy words to describe their query requirements (e.g., student and university). As "Data exploration is about efficiently extracting knowledge from data even if we do not know exactly what we are looking for" [22], visual analysis plays an important role in finding something interesting and generating new questions [24]. Despite the existing methods dedicated to enabling fuzzy query and building interactive environment for examining the trajectories data, two aspects remain underexplored. The first is relevance quantification. When the analysts specify a fuzzy input and acquire a group of trajectories, [zhaosong is here], people ask question such as why the queried trajectories . Similarly, the analysts need to fully understand the semantics of trajectories. [so it is easy to refine the input]. Thus, we should present a efficient and intuitive visual representation to support illustration and interpretation of massive mobile trajectories.

We have developed a visual analytics system. Our system mainly contains components: The to enable the analysts input a fuzzy query input.

[should pipeline and quantification method be separated?] To sum up, our contributions are as follows:

- We introduce A "textualization-and-analysis" pipeline to support efficient text-based query, and relevance quantification between trajectories and user defined query condition.
- We present a multi-faceted visual interface and interaction designs for specifying the query condition, visual depiction and exploration of the massive trajectory.
- we propose two use scenarios based on a real-world human mobile trajectory dataset in a mid-size city of China to demonstrate the efficacy of our approach.